Quality Information

Use Case: I need to provide information about the quality of my data and how it was measured.

Overview

A principle goal of metadata is to ensure that the data they describe can be independently understood and used effectively. Data quality measures and reports play a critical role in achieving this goal. Connecting these to the metadata record is clearly important.

The approach to including quality information in ISO 19115 metadata records is similar to the approach used in ECHO and improves on it considerably. It includes the capability to describe quality measures that are used and the techniques used to apply them. Understanding how to take advantage of this flexibility and implementing systems that maximize the value that this capability provides will certainly be a challenge for the environmental data community.

The Data Quality Section of the ISO standard supports flexibility at several levels. The DQ_Data_Quality object (see Figure) includes two sections: scope, and element. A metadata record can have any number of associated DQ_DataQuality objects.

Presentations available that describe the ISO data quality standard and compare it with the DIF and ECHO approaches.

Recommendations

An important initial goal is to connect users to existing literature that includes quality information. This should be done at the collection level using standaloneQualityReports. A complete report includes a description of the scope of the quality information with spatial and temporal extent information, an abstract, and a complete citation.

Conceptual Model

This Figure shows an overview of the data quality information included in the ISO 19157 standard. The elements are:

Scope: quality information can be provided at many different level of detail from an entire series down to specific attributes in particular regions or time periods. The scope attribute lets users know what the current information pertains to.

StandaloneQualityReport: quality information exists in papers and reports that are already part of the scientific literature. Users can find this information using citations included in the metadata record.

Report: quality reports hold the bulk of the structured quality information in the metadata record. They describe the quality measure, how it was applied, and the result of the application.

An overview of the complete data quality standard is also available.
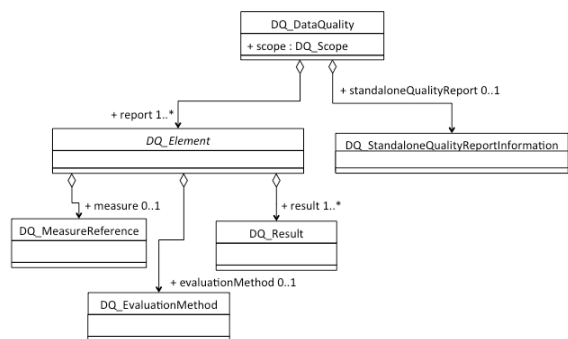
Scope:

The DQ_Scope object enables a complete description of the scope of a data quality report whether it pertains to a complete data series or to a particular temporal and spatial extent for a particular parameter in a dataset.

level:

The level element of the scope provides a general description of the scope using a codelist of standard terms. The level can be used in an initial search of the quality information to answer questions like "what kinds of quality information are available for this collection or granule".

attribute	dimensionGroup	model	aggregate
attributeType	feature	tile	product
collectionHardware	featureType	metadata	collection
collectionSession	propertyType	initiative	coverage
dataset	fieldSession	sample	application
series	software	document	
nonGeographicDataset	service	repository	

levelDescription:

In some cases, the general description provided by the codelist is not enough. In those cases the levelDescription can be used to provide more detail. For example, the scope code might be "attribute", indicating that the report concerns an attribute of the dataset, and the levelDescription would provide the name(s) of the specific attributes covered, i.e. precipitation rate.

extent:

In some cases, the quality information may apply to a particular spatial and temporal extent of the dataset. This extent can be described at several different levels of detail using the EX_Extent object. The extent should be described quantitatively, if possible, or with a text description if quantitative information is not possible.

StandAlone Report

ISO 19157 recognizes that important data quality information can exist outside of the conceptual framework of the model and that it may be helpful to provide that information as a supplement to the metadata. The standaloneQualityReport was added to enable connections between the metadata and these reports. It includes an abstract (CharacterString) that provides a brief description of the quality report in the metadata record and a the citation (CI_Citation) that provides the reference that a user needs for the complete report.
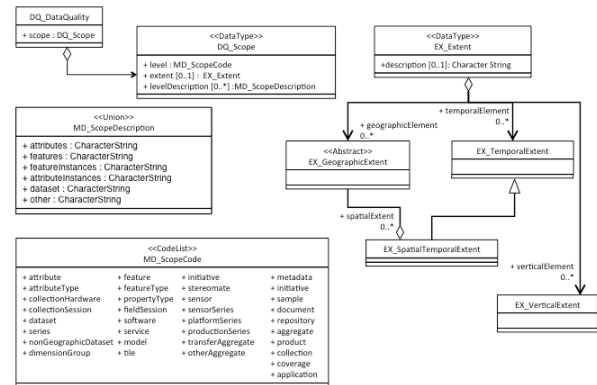
In many cases quality information is provided on the Web pages. Those can also be referenced as StandAloneReports.

As an example, consider this data quality description from a GCMD metadata record:

Abstract: The fire training-set may also have been biased against savanna and savanna woodland fires since their detection is more difficult than in humid, forest environments with cool background temperatures [Malingreau, 1990]. There may, therefore, be an under-sampling of fires in these warmer background environments.

Citation: Malingreau J.P, 1990, The contribution of remote sensing to the global monitoring of fires in tropical and subtropical ecosystems. In: Fire in Tropical Biota, (J.G. Goldammer , editor), Springer Verlag , Berlin: 337-370.

Report

Quality reports are encapsulated in the 19115 DQ_Element class that includes three kinds of information, the measure, the evaluationMethod and the result. It is expected that an organization uses a standard set of quality measures that are managed centrally and have unique names, descriptions, and identifiers. The nameOfMeasure and measureDescription are included in the metadata record and should help users understand the nature of the measure. The measureIdentification is an identifier that can be used to find a more detailed description of the measure.

The evaluation method provides information on how the quality measure was applied for this report. These methods can also be standardized across an organization. Citations are available to describe the specific proceedure and more complete references.

Finally, the result reports the results of the application of the

DQ_MeasureReference [0..1]
+ measureIdentification: MD_Identifier [0..1]
+ nameOfMeasure: CharacterString [0..*]
+ measureDescription: CharacterString [0..1]

DQ_EvaluationMethod [0..1]
+ dateTime: DateTime [0..*]
+ evaluationMethodDescription: CharacterString [0..1]
+ evaluationProceedure: CI_Citation [0..1]
+ referenceDoc: CI_Citation [0..*]
+ evaluationMethodType: DQ_EvaluationMethodTypeCode [0..1]

DQ_Result [1..*]
+ dateTime: DateTime [0..*]
+ resultScope: DQ_ScopeCode [0..1]

measure in this specific case. Several types of reports are included in the standard (see below).

Result Types

The ISO Standard include four types of data quality results in the DQ_Element object.

The first, the DQ_ConformanceResult, describes how the dataset was tested for conformance to a published standard and whether the dataset passed the test.

The second, the DQ_QuantitativeResult, provides a mechanism for describing the results of a quantitative quality evaluation.

Finally, the QE_CoverageResult (from ISO 19115-2), allows a quality result to be expressed as a spatial object. For example, if a gridded dataset has an associated grid of quality flags, that quality grid could be described here. Note also that the coverage results have an associated file and format.

The quality of many datasets varies spatially within the dataset. For example, gridded satellite datasets include grids that provide quality flag values for every pixel. Radiosondes in the atmosphere and profiles in the ocean also can have quality that varies along their paths. In situations like these, the quality information can be described using a spatial feature like a grid or a line. In ISO quality information in spatial features is described using a coverageResult. The coverageResult includes MD_SpatialRepresentation and MD_CoverageDescriptions that describe the quality coverage. In this case, the MD_CoverageDescription can serve a role that is similar to the role of the measure and evaluationMethod elements in the report. The MD_CoverageDescription includes a contentType that can currently be image, thematicRepresentation, or physicalMeasurement. The revision of 19115 may add qualityInformation to this codelist and allow multiple contentType values.

Finally, the DQ_Descriptive result was introduced in ISO 19157 in order to include simple textual descriptions of data quality in the standard. It is a simple CharacterString.

DQ_ConformanceResult
+ specification : CI_Citation
+ explanation : CharacterString
+ pass : Boolean
Or
DQ_QuantitativeResult
+ valueType [0..1] : RecordType
+ valueUnit : UnitOfMeasure
+ errorStatistic [0..1] : CharacterString
+ value [1..*] : Record
Or
QE_CoverageResult (added in ISO 19115-2, see UML below)
+ spatialRepresentationType: MD_SpatialRepresentationTypeCode, one of vector, grid, textTable, tin, stereoModel, video
+ resultFile: MX_DataFile
+ resultFormat: MD_Format
+ resultSpatialRepresentation: MD_SpatialRepresentation
+ resultContentDescription: MD_CoverageDescription
Or
DQ_DescriptiveResult
+ statement: CharacterString

DQ_CoverageResult
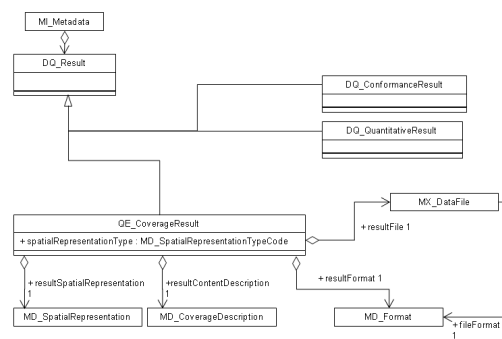
Implementation (ISO)

ISO 19115-2

The general structure of a data quality report in ISO 19115-2 is:

```
<gmd:DQ_DataQuality
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.isotc211.org/2005/gmd
http://www.isotc211.org/2005/gmd/gmd.xsd
http://www.isotc211.org/2005/gmi
http://www.isotc211.org/2005/gmi/gmi.xsd"
```

```xml
    xmlns:gmi="http://www.isotc211.org/2005/gmi"
xmlns:gco="http://www.isotc211.org/2005/gco"
xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:xlink="http://www.w3.org/1999/xlink"
    xmlns:srv="http://www.isotc211.org/2005/srv"
xmlns:gml="http://www.opengis.net/gml/3.2">
    <gmd:scope>
        <!--
        The DQ_Scope object can include an id attribute that allows the
same quality scope to be referenced from several different sections of
the metadata record.
        For example, if all of the quality reports in the record pertain
to the entire granule, they could all reference the same scope object
with a value of "dataset".
        In this case, the id is set to datasetScope.
        -->
        <gmd:DQ_Scope id="datasetScope">
            <gmd:level>
                <!--
                The MD_ScopeCode is a codelist that includes two
standard attributes: a URL that gives the location of the codeList,
                and a value from the codelist. The location can be set
to the standard location if the standard codelist is sufficient
                for the data provider needs or, if necessary, it can be
set to a codelist that has been extended to include dataset
                specific values.
                -->
                <gmd:MD_ScopeCode
codeList="http:CodeListLocation//#MD_ScopeCode"
codeListValue="dataset">dataset</gmd:MD_ScopeCode>
            </gmd:level>
            <!--
            The extent of the quality report can include specific
information, if appropriate, or it can reference an extent defined
            in another part of the metadata record. In this case, the
extent of the quality report is the same as that of the dataset,
            so the identifier boundingExtent is used. If necessary, the
qualityReport can contain more than one extent. See extent examples.
            -->
            <gmd:extent xlink:href="#boundingExtent"/>
            <gmd:levelDescription>
                <!--
                The levelDescription is used to provide more detail
about the scope of the quality report. It can include name(s) of
                attributes or features or a text description of the
scope. The properties also include dataset and other which are
                both character strings. If necessary, the quality report
can contain more than one levelDescription.
                -->
                <gmd:MD_ScopeDescription>
                    <gmd:dataset>

<gco:CharacterString>levelDescription</gco:CharacterString>
```

```
                </gmd:dataset>
              </gmd:MD_ScopeDescription>
            </gmd:levelDescription>
          </gmd:DQ_Scope>
      </gmd:scope>
      <gmd:report>
          <!--
          The report holds the meat of the DQ_DataQuality section of the
metadata. It includes four major sections: the measure,
          the method, the date, and the result. There are a number of
different types of reports that reflect general areas of
          data quality, e.g. DQ_Completeness, DQ_LogicalConsistency,
DQ_PositionalAccuracy, DQ_ThematicAccuracy, DQ_TemporalAccuracy,
          DQ_QuantitativeAttributeAccuracy, and DQ_Usability. These are
very general descriptors that should be used if clearly
          appropriate. At the same time, some of these names were created
with collections of geographic features in mind and so
          may not make sense in the context of other data types. In those
cases, the very general type DQ_QuantitativeAttributeAccuracy
          serves as a generally applicable catch-all.
          -->
          <gmd:DQ_QuantitativeAttributeAccuracy>
              <!--
              Information about the quality measure: To facilitate data
set comparisons, it is helpful if the results in the data quality
              reports are expressed in a comparable way and that there is
a common understanding of the data quality measures that have
              been used. ISO 19157 provides a set of standard measures to
facilitate comparisons, but these are very general, e.g.
              "number of missing items" or "relative horizontal error".
The standard also describes how user defined quality
              measures can be cataloged for sharing across a community.

              A number of types of data quality reports are available. The
DQ_QuantitativeAttributeAccuracy is the most general and
              most broadly applicable of these.

              ISO 19115 quality reports include three elements that are
used to describe the quality measure. The nameOfMeasure and
              measureDescription are used to provide enough information
for a user to understand the measure being used if they are
              familiar with the measures used in a community. The
measureIdentifier is used to support a search for more information.

              -->

              <!--
              ESDIS Data Quality Information:

              The ECHO metadata model includes several kinds of quality
information: QAStats and QAFlags.
              These measures have standard names. The QAStats are
QAPercentMissingData, QAPercentOutOfBoundsData,
```

```
QAPercentInterpolatedData,
            QAPercentCloudCover and the QAFlags are
AutomaticQualityFlag, OperationalQualityFlag, ScienceQualityFlag. In the
ISO
            implementation these are measure names.

            Currently the measure names serve as globally unique
identifiers for the ESDIS quality measures so nameOfMeasure =
identifier.

            The QAStats have self-explanatory names so they do not
require a description. The QAFlags are more general and have mission or
            product specific explanations. In ECHO these are called
AutomaticQualityFlagExplanation, OperationalQualityFlagExplanation,
            ScienceQualityFlagExplanation. In the ISO implementation,
these explanations are measureDescriptions.

            Other quality information:

            Many ESDIS products include quality indicators that are
specific to a particular product and not in the general ECHO
            model. In some cases this quality information is in product
specific attributes (or additional attributes)
            with names that include "PERCENT" or start with "QA", e.g.
PERCENTSHAPEFIXEDBRDFS, PERCENTLOWSUN,
            QAPERCENTPOOROUTPUT250MBAND1, QAPERCENTPOOROUTPUT250MBAND2.
In the ISO implementation these are measure names.

            In some cases these names are self-explanatory, but in many
cases, more detailed description of how they are
            calculated would benefit users. Those descriptions could be
measure descriptions or references to how they
            were calculated could be provided as evaluation procedure
citations.
            -->
            <gmd:nameOfMeasure>
                <gco:CharacterString>nameOfMeasure</gco:CharacterString>
            </gmd:nameOfMeasure>
            <gmd:measureIdentification>
                <gmd:MD_Identifier>
                    <gmd:code>

<gco:CharacterString>measureIdentifier</gco:CharacterString>
                    </gmd:code>
                </gmd:MD_Identifier>
            </gmd:measureIdentification>
            <gmd:measureDescription>

<gco:CharacterString>measureDescription</gco:CharacterString>
            </gmd:measureDescription>
            <!--
                The evaluation method provides a description of how the
measure was used in the context of a particular dataset
```

```
                    in order to generate the quality report. ISO 19115
includes three elements that are relevant to the evaluationMethod:
                    the type is a codelist with values of directInternal,
directExternal, or indirect; the evaluationMethodDescription
                    is a text description that is available to a user along
with the metadata record; the evaluationProceedure provides
                    a citation to more detailed information.
                -->
                <gmd:evaluationMethodType>
                    <gmd:DQ_EvaluationMethodTypeCode codeList="URI to
codelist location"
codeListValue="direct_external">direct_external</gmd:DQ_EvaluationMethod
TypeCode>
                </gmd:evaluationMethodType>
                <gmd:evaluationMethodDescription>

<gco:CharacterString>evaluationMethodDescription</gco:CharacterString>
                </gmd:evaluationMethodDescription>
                <gmd:evaluationProcedure>
                    <gmd:CI_Citation>
                        <gmd:title>
                            <gco:CharacterString>title of reference
material</gco:CharacterString>
                        </gmd:title>
                        <gmd:date gco:nilReason="missing"/>
                    </gmd:CI_Citation>
                </gmd:evaluationProcedure>
                <!--
                    The dateTime gives the date and time of the quality test
in ISO8601 compliant text format
                -->
                <gmd:dateTime>
                    <gco:DateTime>2012-10-02T15:44:00</gco:DateTime>
                </gmd:dateTime>
                <!--
                    The result is the actual result of the quality test.
There are four types of results: DQ_QuantitativeResult,
                    DQ_ConformanceResult, QE_CoverageResult (from ISO
19115-2), and DQ_DescriptiveResult (from ISO 19157). A single
                    report can contain one or two results.
                -->
                <gmd:result>
                    <gmd:DQ_QuantitativeResult>
                        <!--
                            gmd:DQ_QuantitativeResult:

                            Results of many quality reports are single
simple numbers (%'s or counts). In those cases a
                            DQ_QuantitativeResult can be used to hold
the result.

                            Unit definitions generally come from an
accepted source that can be referenced for an authoritative
```

```
                                definition. In Earth Science the UDUNITS
package is commonly used
(http://www.unidata.ucar.edu/software/udunits/).
                                A more general source for unit definitions
is "The Unified Code for Units of Measure" (http://unitsofmeasure.org).
                                See
https://coastwatch.pfeg.noaa.gov/erddap/convert/units.html for a
comparison of the two approaches.

                                The xlink included in the valueUnit is
designed to provide an unambiguous identifier for the definition of the
                                unit. Neither UCUM not UDUNITS support such
URIs. UCUM does have a table of unit abbreviations and definitions
                                at
http://unitsofmeasure.org/ucum.html#section-Alphabetic-Index. These can
be referenced as
                                degrees Celsius =
xlink:href="http://unitsofmeasure.org/ucum.html#Cel"

                            -->
                    <gmd:valueUnit xlink:href="AnyValidURI"/>
                    <gmd:value>
                        <!--
                                A Record is a string of value(s) with a
given type. In most quantitative quality reports the type is a simple
                                primative type (gco:Real_PropertyType,
gco:Integer_PropertyType). In other cases, it could be a string
                                (gco:CharacterString_PropertyType) that
includes multiple numbers.
                            -->
                        <gco:Record xsi:type="gco:Integer_PropertyType">
                            <gco:Integer>999</gco:Integer>
                        </gco:Record>
                    </gmd:value>
                </gmd:DQ_QuantitativeResult>
            </gmd:result>
            <gmd:result>
                <gmi:QE_CoverageResult>
                    <!--
                            gmi:QE_CoverageResult:

                            The gmi:QE_CoverageResult is used to provide
a quality report that is in the form of a coverage (e.g. a
                            timeseries or grid). It therefore shares a
number of elements with coverage descriptions. In cases where
                            these elements are repeated in a single
metadata record, these can be references to elements that are
                            completely described elsewhere in the record
(e.g. xlink:href="#datasetspatialRepresentation").

                            The resultFile and resultFormat can be used
to describe a coverage result that is in a different file.
                        -->
```

```xml
                    <gmi:spatialRepresentationType/>
                    <gmi:resultSpatialRepresentation/>
                    <gmi:resultContentDescription/>
                    <gmi:resultFormat/>
                    <gmi:resultFile/>
                </gmi:QE_CoverageResult>
            </gmd:result>
            <gmd:result>
                <gmd:DQ_ConformanceResult>
                    <!--
                            gmd:DQ_ConformanceResult:

                            The gmd:DQ_ConformanceResult is used to
provide a quality report on conformance of a dataset to a
                            particular standard, e.g. an OGC Web Service
or OPeNDAP.

                            The result in this case is a Boolean, i.e.
True or False.
                        -->
                    <gmd:specification/>
                    <gmd:explanation/>
                    <gmd:pass/>
                </gmd:DQ_ConformanceResult>
            </gmd:result>
        </gmd:DQ_QuantitativeAttributeAccuracy>
    </gmd:report>
</gmd:DQ_DataQuality>
```

ISO 19115-1 / ISO 19157

The current version of the ISO metadata standards (ISO 19157 and ISO 19115-1) improved the structure of the metadata by modularizing the measure and evaluation information. This makes it easier to re-use measure and evaluation information across multiple collections or granules. The namespaces also changed as another step towards increased modularity.

```xml
<mdq:DQ_DataQuality
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://standards.iso.org/iso/19115/-3/mdb/1.0
http://standards.iso.org/iso/19115/-3/mdt/1.0/mdt.xsd"
    xmlns:mdq="http://standards.iso.org/iso/19157/-2/mdq/1.0"
xmlns:mdb="http://standards.iso.org/iso/19115/-3/mdb/1.0"
xmlns:gco="http://standards.iso.org/iso/19115/-3/gco/1.0"
    xmlns:mcc="http://standards.iso.org/iso/19115/-3/mcc/1.0"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:cit="http://standards.iso.org/iso/19115/-3/cit/1.0">
    <mdq:scope>
        <!--
          The DQ_Scope object can include an id attribute that allows the
same quality scope to be referenced from several different sections of
the metadata record.
          For example, if all of the quality reports in the record pertain
to the entire granule, they could all reference the same scope object
with a value of "dataset".
          In this case, the id is set to datasetScope.
          -->
        <mcc:MD_Scope>
          <mcc:level>
              <!--
                  The MD_ScopeCode is a codelist that includes two
standard attributes: a URL that gives the location of the codeList,
                  and a value from the codelist. The location can be set
to the standard location if the standard codelist is sufficient
                  for the data provider needs or, if necessary, it can be
set to a codelist that has been extended to include dataset
                  specific values.
                  -->
              <mcc:MD_ScopeCode
codeList="http:CodeListLocation//#MD_ScopeCode"
codeListValue="dataset">dataset</mcc:MD_ScopeCode>
          </mcc:level>
          <!--
              The extent of the quality report can include specific
information, if appropriate, or it can reference an extent defined
              in another part of the metadata record. In this case, the
extent of the quality report is the same as that of the dataset,
              so the identifier boundingExtent is used. If necessary, the
qualityReport can contain more than one extent. See extent examples.
              -->
```

```xml
            <mcc:extent xlink:href="#boundingExtent"/>
            <mcc:levelDescription>
                <!--
                    The levelDescription is used to provide more detail
about the scope of the quality report. It can include name(s) of
                    attributes or features or a text description of the
scope. The properties also include dataset and other which are
                    both character strings. If necessary, the quality report
can contain more than one levelDescription.
                    -->
                <mcc:MD_ScopeDescription>
                    <mcc:dataset>

<gco:CharacterString>levelDescription</gco:CharacterString>
                    </mcc:dataset>
                </mcc:MD_ScopeDescription>
            </mcc:levelDescription>
        </mcc:MD_Scope>
    </mdq:scope>
    <mdq:standaloneQualityReport>
        <mdq:DQ_StandaloneQualityReportInformation>
            <!--
                DQ_StandaloneQualityReportInformation
                Many important data quality reports are described in
scientific papers or reports that can be referenced as
                standalone quality reports that include citations to the
report and an abstract that provides a brief
                description of the result. This is a new capability
introduced in ISO 19157. The report includes a citation
                -->
            <mdq:reportReference>
                <cit:CI_Citation/>
            </mdq:reportReference>
            <mdq:abstract>
                <gco:CharacterString/>
            </mdq:abstract>
        </mdq:DQ_StandaloneQualityReportInformation>
    </mdq:standaloneQualityReport>
    <mdq:report>
        <!--
            The report holds the meat of the DQ_DataQuality section og the
metadata. It includes four major sections: the measure,
            the method, the date, and the result. There are a number of
different types of reports that reflect general areas of
            data quality, e.g. DQ_Completeness, DQ_LogicalConsistency,
DQ_PositionalAccuracy, DQ_ThematicAccuracy, DQ_TemporalAccuracy,
            DQ_QuantitativeAttributeAccuracy, and DQ_Usability. These are
very general descriptors that should be used if clearly
            appropriate. At the same time, some of these names were created
with collections of geographic features in mind and so
            may not make sense in the context of other data types. In those
cases, the very general type DQ_QuantitativeAttributeAccuracy
            serves as a generally applicable catch-all.
```

```
            -->
        <mdq:DQ_QuantitativeAttributeAccuracy>
           <!--
                Information about the quality measure: To facilitate data
set comparisons, it is helpful if the results in the data quality
                reports are expressed in a comparable way and that there is
a common understanding of the data quality measures that have
                been used. ISO 19157 provides a set of standard measures to
facilitate comparisons, but these are very general, e.g.
                "number of missing items" or "relative horizontal error".
The standard also describes how user defined quality
                measures can be cataloged for sharing across a community.

                A number of types of data quality reports are available. The
DQ_QuantitativeAttributeAccuracy is the most general and
                most broadly applicable of these.

                ISO 19115 quality reports include three elements that are
used to describe the quality measure. The nameOfMeasure and
                measureDescription are used to provide enough information
for a user to understand the measure being used if they are
                familiar with the measures used in a community. The
measureIdentifier is used to support a search for more information.

                -->

            <!--
                ESDIS Data Quality Information:

                The ECHO metadata model includes several kinds of quality
information: QAStats and QAFlags.
                These measures have standard names. The QAStats are
QAPercentMissingData, QAPercentOutOfBoundsData,
QAPercentInterpolatedData,
                QAPercentCloudCover and the QAFlags are
AutomaticQualityFlag, OperationalQualityFlag, ScienceQualityFlag. In the
ISO
                implementation these are measure names.

                Currently the measure names serve as globally unique
identifiers for the ESDIS quality measures so nameOfMeasure =
identifier.

                The QAStats have self-explanatory names so they do not
require a description. The QAFlags are more general and have mission or
                product specific explanations. In ECHO these are called
AutomaticQualityFlagExplanation, OperationalQualityFlagExplanation,
                ScienceQualityFlagExplanation. In the ISO implementation,
these explanations are measureDescriptions.

                Other quality information:

                Many ESDIS products include quality indicators that are
```

specific to a particular product and not in the general ECHO
            model. In some cases this quality information is in product
specific attributes (or additional attributes)
            with names that include "PERCENT" or start with "QA", e.g.
PERCENTSHAPEFIXEDBRDFS, PERCENTLOWSUN,
            QAPERCENTPOOROUTPUT250MBAND1, QAPERCENTPOOROUTPUT250MBAND2.
In the ISO implementation these are measure names.

            In some cases these names are self-explanatory, but in many
cases, more detailed description of how they are
            calculated would benefit users. Those descriptions could be
measure descriptions or references to how they
            were calculated could be provided as evaluation procedure
citations.
            -->
        <mdq:measure>
            <mdq:DQ_MeasureReference>
                <mdq:measureIdentification>
                    <mcc:MD_Identifier>
                        <mcc:code>

<gco:CharacterString>measureIdentifier</gco:CharacterString>
                        </mcc:code>
                    </mcc:MD_Identifier>
                </mdq:measureIdentification>
                <mdq:nameOfMeasure>

<gco:CharacterString>nameOfMeasure</gco:CharacterString>
                </mdq:nameOfMeasure>
                <mdq:measureDescription>

<gco:CharacterString>measureDescription</gco:CharacterString>
                </mdq:measureDescription>
            </mdq:DQ_MeasureReference>
        </mdq:measure>
        <!--
                The evaluation method provides a description of how the
measure was used in the context of a particular dataset
                in order to generate the quality report. ISO 19157
includes a new abstract DQ_DataEvaluation class that can be implemented
                as DQ_FullInspection, DQ_SampleBasedInspection, or
DQ_IndirectEvaluation. It includes three elements that are relevant to
the evaluationMethod:
                the type is a codelist with values of directInternal,
directExternal, or indirect; the evaluationMethodDescription
                is a text description that is available to a user along
with the metadata record; the evaluationProceedure provides
                a citation to more detailed information.
            -->
        <mdq:evaluationMethod>
            <mdq:DQ_FullInspection>
                <!--
                        The dateTime gives the date and time of the

```xml
quality test in ISO1601 compliant text format
                -->
                <mdq:dateTime>
                    <gco:DateTime>2012-10-02T15:44:00</gco:DateTime>
                </mdq:dateTime>
                <mdq:evaluationMethodDescription>

<gco:CharacterString>evaluationMethodDescription</gco:CharacterString>
                </mdq:evaluationMethodDescription>
                <mdq:evaluationProcedure>
                    <cit:CI_Citation>
                        <cit:title>
                            <gco:CharacterString>title of reference
material</gco:CharacterString>
                        </cit:title>
                        <cit:date gco:nilReason="missing"/>
                    </cit:CI_Citation>
                </mdq:evaluationProcedure>
                <mdq:evaluationMethodType>
                    <mdq:DQ_EvaluationMethodTypeCode
codeList="codeListLocation#DQ_EvaluationMethodTypeCode"
codeListValue="direct_external">direct_external</mdq:DQ_EvaluationMethod
TypeCode>
                </mdq:evaluationMethodType>
            </mdq:DQ_FullInspection>
        </mdq:evaluationMethod>
        <!--
                The result is the actual result of the quality test.
There are four types of results: DQ_QuantitativeResult,
                DQ_ConformanceResult, QE_CoverageResult (from ISO
19115-2), and DQ_DescriptiveResult (from ISO 19157). A single
                report can contain one or two results.
            -->
        <mdq:result>
            <mdq:DQ_QuantitativeResult>
                <!--
                        gmd:DQ_QuantitativeResult:

                        Results of many quality reports are single
simple numbers (%'s or counts). In those cases a
                        DQ_QuantitativeResult can be used to hold
the result.

                        Unit definitions generally come from an
accepted source that can be referenced for an authoritative
                        definition. In Earth Science the UDUNITS
package is commonly used
(http://www.unidata.ucar.edu/software/udunits/).
                        A more general source for unit definitions
is "The Unified Code for Units of Measure" (http://unitsofmeasure.org).
                        See
https://coastwatch.pfeg.noaa.gov/erddap/convert/units.html for a
comparison of the two approaches.
```

```
                                The xlink included in the valueUnit is
designed to provide an unambiguous identifier for the definition of the
                                unit. Neither UCUM not UDUNITS support such
URIs. UCUM does have a table of unit abbreviations and definitions
                                at
http://unitsofmeasure.org/ucum.html#section-Alphabetic-Index. These can
be referenced as
                                degrees Celsius =
xlink:href="http://unitsofmeasure.org/ucum.html#Cel"

                            -->
                <mdq:value>
                    <!--
                                    A Record is a string of value(s) with a
given type. In most quantitative quality reports the type is a simple
                                    primative type (gco:Real_PropertyType,
gco:Integer_PropertyType). In other cases, it could be a string
                                    (gco:CharacterString_PropertyType) that
includes multiple numbers.
                                -->
                    <gco:Record xsi:type="gco:Integer_PropertyType">
                        <gco:Integer>999</gco:Integer>
                    </gco:Record>
                </mdq:value>
                <mdq:valueUnit xlink:href="AnyValidURI"/>
            </mdq:DQ_QuantitativeResult>
        </mdq:result>
        <mdq:result>
            <mdq:QE_CoverageResult>
                <!--
                            gmi:QE_CoverageResult:

                            The gmi:QE_CoverageResult is used to provide
a quality report that is in the form of a coverage (e.g. a
                            timeseries or grid). It therefore shares a
number of elements with coverage descriptions. In cases where
                            these elements are repeated in a single
metadata record, these can be references to elements that are
                            completely described elsewhere in the record
(e.g. xlink:href="#datasetspatialRepresentation").

                            The resultFile and resultFormat can be used
to describe a coverage result that is in a different file.
                        -->
                <mdq:spatialRepresentationType/>
                <mdq:resultSpatialRepresentation/>
                <mdq:resultContentDescription/>
                <mdq:resultFormat/>
                <mdq:resultFile/>
            </mdq:QE_CoverageResult>
        </mdq:result>
        <mdq:result>
```

```xml
            <mdq:DQ_ConformanceResult>
                <!--
                            gmd:DQ_ConformanceResult:

                            The gmd:DQ_ConformanceResult is used to
provide a quality report on conformance of a dataset to a
                            particular standard, e.g. an OGC Web Service
or OPeNDAP.

                            The result in this case is a Boolean, i.e.
True or False.
                    -->
                <mdq:specification/>
                <mdq:explanation/>
                <mdq:pass/>
            </mdq:DQ_ConformanceResult>
        </mdq:result>
        <mdq:result>
            <mdq:DQ_DescriptiveResult>
                <!--
                    DQ_DescriptiveResult
                    The DQ_DescriptiveResult provides an opportunity to
include a character string describing the quality test.
                    These are similar to the dataQuality element in the
DIF metadata standard.
                -->
                <mdq:statement/>
            </mdq:DQ_DescriptiveResult>
        </mdq:result>
    </mdq:DQ_QuantitativeAttributeAccuracy>
```

```
        </mdq:report>
    </mdq:DQ_DataQuality>
```

ECHO Reports in ISO

Details and examples of how ECHO data quality content can be represented in ISO.

Implementation (NcML)

The XML implementations shown above work well in systems designed to manage metadata in internationally accepted XML dialects. Typically these systems manage metadata for collections and, in many cases, they are separate from the actual data files (or granules). Historically, the metadata in files has been generated and used by analysis and visualization tools. It is written and read by tools that understand the specifics of the data format and the organization and naming of the metadata elements. The ISO standards shown above provide an organization and a set of names. Can we take advantage of those in the granules?

We have recently explored this idea with a standard transform between ISO metadata and NcML, a simple XML dialect originally developed as part of the netCDF ecosystem that includes groups and attributes and translated easily to the structure of an HDF file. The NcML version of the ISO 19115 quality information is shown below. This representation includes all of the ISO content and can be transformed back into ISO when it is appropriate.

```
<nc:netcdf xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="http://www.unidata.ucar.edu/namespaces/netcdf/ncml-2.2"

xmlns:nc="http://www.unidata.ucar.edu/namespaces/netcdf/ncml-2.2">
    <nc:group name="gmd:DQ_DataQuality">
        <nc:attribute name="__role" value="metadataRoot"/>
        <nc:attribute name="__type" value="gmd:DQ_DataQuality"/>
        <nc:attribute name="@xsi:schemaLocation"
                      value="http://www.isotc211.org/2005/gmd
http://www.isotc211.org/2005/gmd/gmd.xsd
http://www.isotc211.org/2005/gmi
http://www.isotc211.org/2005/gmi/gmi.xsd"/>
        <!--
          The DQ_Scope object can include an id attribute that allows the
same quality scope to be referenced from several different sections of
the metadata record.
          For example, if all of the quality reports in the record pertain
to the entire granule, they could all reference the same scope object
with a value of "dataset".
          In this case, the id is set to datasetScope.
          -->
        <nc:group name="gmd:scope">
            <nc:attribute name="__role" value="gmd:scope"/>
            <nc:attribute name="__type" value="gmd:DQ_Scope"/>
            <nc:attribute name="@id" value="datasetScope"/>
            <!--
                The MD_ScopeCode is a codelist that includes two
standard attributes: a URL that gives the location of the codeList,
                and a value from the codelist. The location can be set
to the standard location if the standard codelist is sufficient
                for the data provider needs or, if necessary, it can be
```

```
            set to a codelist that has been extended to include dataset
                        specific values.
                        -->
            <nc:group name="gmd:level">
                <nc:attribute name="__role" value="gmd:level"/>
                <nc:attribute name="__type" value="gmd:MD_ScopeCode"/>
                <nc:attribute name="@codeList"
value="http:CodeListLocation//#MD_ScopeCode"/>
                <nc:attribute name="@codeListValue" value="dataset"/>
                <nc:attribute name="value" value="dataset"/>
            </nc:group>
            <!--
                The extent of the quality report can include specific
information, if appropriate, or it can reference an extent defined
                in another part of the metadata record. In this case, the
extent of the quality report is the same as that of the dataset,
                so the identifier boundingExtent is used. If necessary, the
qualityReport can contain more than one extent. See extent examples.
                -->
            <nc:group name="gmd:extent">
                <nc:attribute name="@xlink:href" value="#boundingExtent"/>
            </nc:group>
            <!--
                    The levelDescription is used to provide more detail
about the scope of the quality report. It can include name(s) of
                    attributes or features or a text description of the
scope. The properties also include dataset and other which are
                    both character strings. If necessary, the quality report
can contain more than one levelDescription.
                    -->
            <nc:group name="gmd:levelDescription">
                <nc:attribute name="__role" value="gmd:levelDescription"/>
                <nc:attribute name="__type"
value="gmd:MD_ScopeDescription"/>
                <nc:attribute name="gmd:dataset"
                              value="levelDescription"
                              __isoType="gco:CharacterString"/>
            </nc:group>
        </nc:group>
        <!--
            The report holds the meat of the DQ_DataQuality section of the
metadata. It includes four major sections: the measure,
            the method, the date, and the result. There are a number of
different types of reports that reflect general areas of
            data quality, e.g. DQ_Completeness, DQ_LogicalConsistency,
DQ_PositionalAccuracy, DQ_ThematicAccuracy, DQ_TemporalAccuracy,
            DQ_QuantitativeAttributeAccuracy, and DQ_Usability. These are
very general descriptors that should be used if clearly
            appropriate. At the same time, some of these names were created
with collections of geographic features in mind and so
            may not make sense in the context of other data types. In those
cases, the very general type DQ_QuantitativeAttributeAccuracy
            serves as a generally applicable catch-all.
```

```
        -->
    <nc:group name="gmd:report">
        <nc:attribute name="__role" value="gmd:report"/>
        <nc:attribute name="__type"
value="gmd:DQ_QuantitativeAttributeAccuracy"/>
        <!--
            Information about the quality measure: To facilitate data
set comparisons, it is helpful if the results in the data quality
            reports are expressed in a comparable way and that there is
a common understanding of the data quality measures that have
            been used. ISO 19157 provides a set of standard measures to
facilitate comparisons, but these are very general, e.g.
            "number of missing items" or "relative horizontal error".
The standard also describes how user defined quality
            measures can be cataloged for sharing across a community.

            A number of types of data quality reports are available. The
DQ_QuantitativeAttributeAccuracy is the most general and
            most broadly applicable of these.

            ISO 19115 quality reports include three elements that are
used to describe the quality measure. The nameOfMeasure and
            measureDescription are used to provide enough information
for a user to understand the measure being used if they are
            familiar with the measures used in a community. The
measureIdentifier is used to support a search for more information.

            -->
        <!--
            ESDIS Data Quality Information:

            The ECHO metadata model includes several kinds of quality
information: QAStats and QAFlags.
            These measures have standard names. The QAStats are
QAPercentMissingData, QAPercentOutOfBoundsData,
QAPercentInterpolatedData,
            QAPercentCloudCover and the QAFlags are
AutomaticQualityFlag, OperationalQualityFlag, ScienceQualityFlag. In the
ISO
            implementation these are measure names.

            Currently the measure names serve as globally unique
identifiers for the ESDIS quality measures so nameOfMeasure =
identifier.

            The QAStats have self-explanatory names so they do not
require a description. The QAFlags are more general and have mission or
            product specific explanations. In ECHO these are called
AutomaticQualityFlagExplanation, OperationalQualityFlagExplanation,
            ScienceQualityFlagExplanation. In the ISO implementation,
these explanations are measureDescriptions.

            Other quality information:
```

```
            Many ESDIS products include quality indicators that are
specific to a particular product and not in the general ECHO
            model. In some cases this quality information is in product
specific attributes (or additional attributes)
            with names that include "PERCENT" or start with "QA", e.g.
PERCENTSHAPEFIXEDBRDFS, PERCENTLOWSUN,
            QAPERCENTPOOROUTPUT250MBAND1, QAPERCENTPOOROUTPUT250MBAND2.
In the ISO implementation these are measure names.

            In some cases these names are self-explanatory, but in many
cases, more detailed description of how they are
            calculated would benefit users. Those descriptions could be
measure descriptions or references to how they
            were calculated could be provided as evaluation procedure
citations.
            -->
        <nc:attribute name="gmd:nameOfMeasure"
                      value="nameOfMeasure"
                      __isoType="gco:CharacterString"/>
        <nc:group name="gmd:measureIdentification">
            <nc:attribute name="__role"
value="gmd:measureIdentification"/>
            <nc:attribute name="__type" value="gmd:MD_Identifier"/>
            <nc:attribute name="gmd:code"
                          value="measureIdentifier"
                          __isoType="gco:CharacterString"/>
        </nc:group>
        <nc:attribute name="gmd:measureDescription"
                      value="measureDescription"
                      __isoType="gco:CharacterString"/>
        <!--
                The evaluation method provides a description of how the
measure was used in the context of a particular dataset
                in order to generate the quality report. ISO 19115
includes three elements that are relevant to the evaluationMethod:
                the type is a codelist with values of directInternal,
directExternal, or indirect; the evaluationMethodDescription
                is a text description that is available to a user along
with the metadata record; the evaluationProceedure provides
                a citation to more detailed information.
            -->
        <nc:group name="gmd:evaluationMethodType">
            <nc:attribute name="__role"
value="gmd:evaluationMethodType"/>
            <nc:attribute name="__type"
value="gmd:DQ_EvaluationMethodTypeCode"/>
            <nc:attribute name="@codeList" value="URI to codelist
location"/>
            <nc:attribute name="@codeListValue"
value="direct_external"/>
            <nc:attribute name="value" value="direct_external"/>
        </nc:group>
```

```xml
            <nc:attribute name="gmd:evaluationMethodDescription"
                          value="evaluationMethodDescription"
                          __isoType="gco:CharacterString"/>
            <nc:group name="gmd:evaluationProcedure">
                <nc:attribute name="__role"
value="gmd:evaluationProcedure"/>
                <nc:attribute name="__type" value="gmd:CI_Citation"/>
                <nc:attribute name="gmd:title"
                              value="title of reference material"
                              __isoType="gco:CharacterString"/>
                <nc:group name="gmd:date">
                    <nc:attribute name="@gco:nilReason" value="missing"/>
                </nc:group>
            </nc:group>
            <!--
                   The dateTime gives the date and time of the quality test
in ISO8601 compliant text format
                -->
            <nc:attribute name="gmd:dateTime"
                          value="2012-10-02T15:44:00"
                          __isoType="gco:DateTime"/>
            <!--
                   The result is the actual result of the quality test.
There are four types of results: DQ_QuantitativeResult,
                   DQ_ConformanceResult, QE_CoverageResult (from ISO
19115-2), and DQ_DescriptiveResult (from ISO 19157). A single
                   report can contain one or two results.
                -->
            <nc:group name="gmd:result">
                <nc:attribute name="__role" value="gmd:result"/>
                <nc:attribute name="__type"
value="gmd:DQ_QuantitativeResult"/>
                <!--
                            gmd:DQ_QuantitativeResult:

                            Results of many quality reports are single
simple numbers (%'s or counts). In those cases a
                            DQ_QuantitativeResult can be used to hold
the result.

                            Unit definitions generally come from an
accepted source that can be referenced for an authoritative
                            definition. In Earth Science the UDUNITS
package is commonly used
(http://www.unidata.ucar.edu/software/udunits/).
                            A more general source for unit definitions
is "The Unified Code for Units of Measure" (http://unitsofmeasure.org).
                            See
https://coastwatch.pfeg.noaa.gov/erddap/convert/units.html for a
comparison of the two approaches.

                            The xlink included in the valueUnit is
designed to provide an unambiguous identifier for the definition of the
```

```
                               unit. Neither UCUM not UDUNITS support such
URIs. UCUM does have a table of unit abbreviations and definitions
                               at
http://unitsofmeasure.org/ucum.html#section-Alphabetic-Index. These can
be referenced as
                               degrees Celsius =
xlink:href="http://unitsofmeasure.org/ucum.html#Cel"

                         -->
            <nc:group name="gmd:valueUnit">
                <nc:attribute name="@xlink:href" value="AnyValidURI"/>
            </nc:group>
            <!--
                               A Record is a string of value(s) with a
given type. In most quantitative quality reports the type is a simple
                               primative type (gco:Real_PropertyType,
gco:Integer_PropertyType). In other cases, it could be a string
                               (gco:CharacterString_PropertyType) that
includes multiple numbers.
                         -->
            <nc:group name="gmd:value">
                <nc:attribute name="__role" value="gmd:value"/>
                <nc:attribute name="__type" value="gco:Record"/>
                <nc:attribute name="@xsi:type"
value="gco:Integer_PropertyType"/>
                <nc:attribute name="gco:Record" value="999"
__isoType="gco:Integer"/>
            </nc:group>
        </nc:group>
        <nc:group name="gmd:result_d1e62">
            <nc:attribute name="__role" value="gmd:result"/>
            <nc:attribute name="__type" value="gmi:QE_CoverageResult"/>
            <!--
                               gmi:QE_CoverageResult:

                               The gmi:QE_CoverageResult is used to provide
a quality report that is in the form of a coverage (e.g. a
                               timeseries or grid). It therefore shares a
number of elements with coverage descriptions. In cases where
                               these elements are repeated in a single
metadata record, these can be references to elements that are
                               completely described elsewhere in the record
(e.g. xlink:href="#datasetspatialRepresentation").

                               The resultFile and resultFormat can be used
to describe a coverage result that is in a different file.
                         -->
            <nc:group name="gmi:spatialRepresentationType">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="gmi:resultSpatialRepresentation">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
```

```xml
            <nc:group name="gmi:resultContentDescription">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="gmi:resultFormat">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="gmi:resultFile">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
        </nc:group>
        <nc:group name="gmd:result_d1e70">
            <nc:attribute name="__role" value="gmd:result"/>
            <nc:attribute name="__type"
value="gmd:DQ_ConformanceResult"/>
            <!--
                            gmd:DQ_ConformanceResult:

                            The gmd:DQ_ConformanceResult is used to
provide a quality report on conformance of a dataset to a
                            particular standard, e.g. an OGC Web Service
or OPeNDAP.

                            The result in this case is a Boolean, i.e.
True or False.
                        -->
            <nc:group name="gmd:specification">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="gmd:explanation">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="gmd:pass">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
        </nc:group>
    </nc:group>
    <nc:group name="namespace_1">
        <nc:attribute name="prefix" value="xml"/>
        <nc:attribute name="uri"
value="http://www.w3.org/XML/1998/namespace"/>
    </nc:group>
    <nc:group name="namespace_2">
        <nc:attribute name="prefix" value="xsi"/>
        <nc:attribute name="uri"
value="http://www.w3.org/2001/XMLSchema-instance"/>
    </nc:group>
    <nc:group name="namespace_3">
        <nc:attribute name="prefix" value="gmi"/>
        <nc:attribute name="uri"
value="http://www.isotc211.org/2005/gmi"/>
    </nc:group>
    <nc:group name="namespace_4">
        <nc:attribute name="prefix" value="gco"/>
```

```xml
            <nc:attribute name="uri"
value="http://www.isotc211.org/2005/gco"/>
        </nc:group>
        <nc:group name="namespace_5">
            <nc:attribute name="prefix" value="gmd"/>
            <nc:attribute name="uri"
value="http://www.isotc211.org/2005/gmd"/>
        </nc:group>
        <nc:group name="namespace_6">
            <nc:attribute name="prefix" value="xlink"/>
            <nc:attribute name="uri" value="http://www.w3.org/1999/xlink"/>
        </nc:group>
        <nc:group name="namespace_7">
            <nc:attribute name="prefix" value="srv"/>
            <nc:attribute name="uri"
value="http://www.isotc211.org/2005/srv"/>
        </nc:group>
        <nc:group name="namespace_8">
            <nc:attribute name="prefix" value="gml"/>
            <nc:attribute name="uri"
value="http://www.opengis.net/gml/3.2"/>
        </nc:group>
```

```
        </nc:group>
    </nc:netcdf>
```

ISO 19115-1 / ISO 19157

The same transform can be used to create NcML from the ISO 19115-1 / ISO 19157 representation.

```
<nc:netcdf xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xmlns="http://www.unidata.ucar.edu/namespaces/netcdf/ncml-2.2"

xmlns:nc="http://www.unidata.ucar.edu/namespaces/netcdf/ncml-2.2">
    <nc:group name="mdq:DQ_DataQuality">
        <nc:attribute name="__role" value="metadataRoot"/>
        <nc:attribute name="__type" value="mdq:DQ_DataQuality"/>
        <nc:attribute name="@xsi:schemaLocation"
                      value="http://standards.iso.org/iso/19115/-3/mdb/1.0
http://standards.iso.org/iso/19115/-3/mdt/1.0/mdt.xsd"/>
        <!--
          The DQ_Scope object can include an id attribute that allows the
same quality scope to be referenced from several different sections of
the metadata record.
          For example, if all of the quality reports in the record pertain
to the entire granule, they could all reference the same scope object
with a value of "dataset".
          In this case, the id is set to datasetScope.
          -->
        <nc:group name="mdq:scope">
          <nc:attribute name="__role" value="mdq:scope"/>
          <nc:attribute name="__type" value="mcc:MD_Scope"/>
          <!--
                  The MD_ScopeCode is a codelist that includes two
standard attributes: a URL that gives the location of the codeList,
                  and a value from the codelist. The location can be set
to the standard location if the standard codelist is sufficient
                  for the data provider needs or, if necessary, it can be
set to a codelist that has been extended to include dataset
                  specific values.
                  -->
          <nc:group name="mcc:level">
              <nc:attribute name="__role" value="mcc:level"/>
              <nc:attribute name="__type" value="mcc:MD_ScopeCode"/>
              <nc:attribute name="@codeList"
value="http:CodeListLocation//#MD_ScopeCode"/>
              <nc:attribute name="@codeListValue" value="dataset"/>
              <nc:attribute name="value" value="dataset"/>
          </nc:group>
          <!--
              The extent of the quality report can include specific
information, if appropriate, or it can reference an extent defined
```

```xml
                  in another part of the metadata record. In this case, the
extent of the quality report is the same as that of the dataset,
                  so the identifier boundingExtent is used. If necessary, the
qualityReport can contain more than one extent. See extent examples.
                  -->
            <nc:group name="mcc:extent">
                <nc:attribute name="@xlink:href" value="#boundingExtent"/>
            </nc:group>
            <!--
                      The levelDescription is used to provide more detail
about the scope of the quality report. It can include name(s) of
                      attributes or features or a text description of the
scope. The properties also include dataset and other which are
                      both character strings. If necessary, the quality report
can contain more than one levelDescription.
                      -->
            <nc:group name="mcc:levelDescription">
                <nc:attribute name="__role" value="mcc:levelDescription"/>
                <nc:attribute name="__type"
value="mcc:MD_ScopeDescription"/>
                <nc:attribute name="mcc:dataset"
                              value="levelDescription"
                              __isoType="gco:CharacterString"/>
            </nc:group>
        </nc:group>
        <nc:group name="mdq:standaloneQualityReport">
            <nc:attribute name="__role"
value="mdq:standaloneQualityReport"/>
            <nc:attribute name="__type"
value="mdq:DQ_StandaloneQualityReportInformation"/>
            <!--
                DQ_StandaloneQualityReportInformation
                Many important data quality reports are described in
scientific papers or reports that can be referenced as
                standalone quality reports that include citations to the
report and an abstract that provides a brief
                description of the result. This is a new capability
introduced in ISO 19157. The report includes a citation
            -->
            <nc:group name="mdq:reportReference">
                <nc:attribute name="__role" value="mdq:reportReference"/>
                <nc:attribute name="__type" value="cit:CI_Citation"/>
            </nc:group>
            <nc:group name="gco:CharacterString">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
        </nc:group>
        <!--
            The report holds the meat of the DQ_DataQuality section og the
metadata. It includes four major sections: the measure,
            the method, the date, and the result. There are a number of
different types of reports that reflect general areas of
            data quality, e.g. DQ_Completeness, DQ_LogicalConsistency,
```

```
DQ_PositionalAccuracy, DQ_ThematicAccuracy, DQ_TemporalAccuracy,
        DQ_QuantitativeAttributeAccuracy, and DQ_Usability. These are
very general descriptors that should be used if clearly
        appropriate. At the same time, some of these names were created
with collections of geographic features in mind and so
        may not make sense in the context of other data types. In those
cases, the very general type DQ_QuantitativeAttributeAccuracy
        serves as a generally applicable catch-all.
        -->
    <nc:group name="mdq:report">
        <nc:attribute name="__role" value="mdq:report"/>
        <nc:attribute name="__type"
value="mdq:DQ_QuantitativeAttributeAccuracy"/>
        <!--
            Information about the quality measure: To facilitate data
set comparisons, it is helpful if the results in the data quality
            reports are expressed in a comparable way and that there is
a common understanding of the data quality measures that have
            been used. ISO 19157 provides a set of standard measures to
facilitate comparisons, but these are very general, e.g.
            "number of missing items" or "relative horizontal error".
The standard also describes how user defined quality
            measures can be cataloged for sharing across a community.

            A number of types of data quality reports are available. The
DQ_QuantitativeAttributeAccuracy is the most general and
            most broadly applicable of these.

            ISO 19115 quality reports include three elements that are
used to describe the quality measure. The nameOfMeasure and
            measureDescription are used to provide enough information
for a user to understand the measure being used if they are
            familiar with the measures used in a community. The
measureIdentifier is used to support a search for more information.

            -->
        <!--
            ESDIS Data Quality Information:

            The ECHO metadata model includes several kinds of quality
information: QAStats and QAFlags.
            These measures have standard names. The QAStats are
QAPercentMissingData, QAPercentOutOfBoundsData,
QAPercentInterpolatedData,
            QAPercentCloudCover and the QAFlags are
AutomaticQualityFlag, OperationalQualityFlag, ScienceQualityFlag. In the
ISO
            implementation these are measure names.

            Currently the measure names serve as globally unique
identifiers for the ESDIS quality measures so nameOfMeasure =
identifier.
```

The QAStats have self-explanatory names so they do not
require a description. The QAFlags are more general and have mission or
            product specific explanations. In ECHO these are called
AutomaticQualityFlagExplanation, OperationalQualityFlagExplanation,
            ScienceQualityFlagExplanation. In the ISO implementation,
these explanations are measureDescriptions.

            Other quality information:

            Many ESDIS products include quality indicators that are
specific to a particular product and not in the general ECHO
            model. In some cases this quality information is in product
specific attributes (or additional attributes)
            with names that include "PERCENT" or start with "QA", e.g.
PERCENTSHAPEFIXEDBRDFS, PERCENTLOWSUN,
            QAPERCENTPOOROUTPUT250MBAND1, QAPERCENTPOOROUTPUT250MBAND2.
In the ISO implementation these are measure names.

            In some cases these names are self-explanatory, but in many
cases, more detailed description of how they are
            calculated would benefit users. Those descriptions could be
measure descriptions or references to how they
            were calculated could be provided as evaluation procedure
citations.
            -->
        <nc:group name="mdq:measure">
            <nc:attribute name="__role" value="mdq:measure"/>
            <nc:attribute name="__type"
value="mdq:DQ_MeasureReference"/>
            <nc:group name="mdq:measureIdentification">
                <nc:attribute name="__role"
value="mdq:measureIdentification"/>
                <nc:attribute name="__type" value="mcc:MD_Identifier"/>
                <nc:attribute name="mcc:code"
                        value="measureIdentifier"
                        __isoType="gco:CharacterString"/>
            </nc:group>
            <nc:attribute name="mdq:nameOfMeasure"
                        value="nameOfMeasure"
                        __isoType="gco:CharacterString"/>
            <nc:attribute name="mdq:measureDescription"
                        value="measureDescription"
                        __isoType="gco:CharacterString"/>
        </nc:group>
        <!--
                The evaluation method provides a description of how the
measure was used in the context of a particular dataset
                in order to generate the quality report. ISO 19157
includes a new abstract DQ_DataEvaluation class that can be implemented
                as DQ_FullInspection, DQ_SampleBasedInspection, or
DQ_IndirectEvaluation. It includes three elements that are relevant to
the evaluationMethod:
                the type is a codelist with values of directInternal,

```
directExternal, or indirect; the evaluationMethodDescription
               is a text description that is available to a user along
with the metadata record; the evaluationProceedure provides
               a citation to more detailed information.
            -->
        <nc:group name="mdq:evaluationMethod">
            <nc:attribute name="__role" value="mdq:evaluationMethod"/>
            <nc:attribute name="__type" value="mdq:DQ_FullInspection"/>
            <!--
                    The dateTime gives the date and time of the
quality test in ISO1601 compliant text format
                -->
            <nc:attribute name="mdq:dateTime"
                        value="2012-10-02T15:44:00"
                        __isoType="gco:DateTime"/>
            <nc:attribute name="mdq:evaluationMethodDescription"
                        value="evaluationMethodDescription"
                        __isoType="gco:CharacterString"/>
            <nc:group name="mdq:evaluationProcedure">
                <nc:attribute name="__role"
value="mdq:evaluationProcedure"/>
                <nc:attribute name="__type" value="cit:CI_Citation"/>
                <nc:attribute name="cit:title"
                            value="title of reference material"
                            __isoType="gco:CharacterString"/>
                <nc:group name="cit:date">
                    <nc:attribute name="@gco:nilReason" value="missing"/>
                </nc:group>
            </nc:group>
            <nc:group name="mdq:evaluationMethodType">
                <nc:attribute name="__role"
value="mdq:evaluationMethodType"/>
                <nc:attribute name="__type"
value="mdq:DQ_EvaluationMethodTypeCode"/>
                <nc:attribute name="@codeList"
value="codeListLocation#DQ_EvaluationMethodTypeCode"/>
                <nc:attribute name="@codeListValue"
value="direct_external"/>
                <nc:attribute name="value" value="direct_external"/>
            </nc:group>
        </nc:group>
        <!--
                The result is the actual result of the quality test.
There are four types of results: DQ_QuantitativeResult,
                DQ_ConformanceResult, QE_CoverageResult (from ISO
19115-2), and DQ_DescriptiveResult (from ISO 19157). A single
                report can contain one or two results.
            -->
        <nc:group name="mdq:result">
            <nc:attribute name="__role" value="mdq:result"/>
            <nc:attribute name="__type"
value="mdq:DQ_QuantitativeResult"/>
            <!--
```

```
                             gmd:DQ_QuantitativeResult:

                             Results of many quality reports are single
simple numbers (%'s or counts). In those cases a
                             DQ_QuantitativeResult can be used to hold
the result.

                             Unit definitions generally come from an
accepted source that can be referenced for an authoritative
                             definition. In Earth Science the UDUNITS
package is commonly used
(http://www.unidata.ucar.edu/software/udunits/).
                             A more general source for unit definitions
is "The Unified Code for Units of Measure" (http://unitsofmeasure.org).
                             See
https://coastwatch.pfeg.noaa.gov/erddap/convert/units.html for a
comparison of the two approaches.

                             The xlink included in the valueUnit is
designed to provide an unambiguous identifier for the definition of the
                             unit. Neither UCUM not UDUNITS support such
URIs. UCUM does have a table of unit abbreviations and definitions
                             at
http://unitsofmeasure.org/ucum.html#section-Alphabetic-Index. These can
be referenced as
                             degrees Celsius =
xlink:href="http://unitsofmeasure.org/ucum.html#Cel"

                         -->
            <!--
                             A Record is a string of value(s) with a
given type. In most quantitative quality reports the type is a simple
                             primative type (gco:Real_PropertyType,
gco:Integer_PropertyType). In other cases, it could be a string
                             (gco:CharacterString_PropertyType) that
includes multiple numbers.
                         -->
            <nc:group name="mdq:value">
                <nc:attribute name="__role" value="mdq:value"/>
                <nc:attribute name="__type" value="gco:Record"/>
                <nc:attribute name="@xsi:type"
value="gco:Integer_PropertyType"/>
                <nc:attribute name="gco:Record" value="999"
__isoType="gco:Integer"/>
            </nc:group>
            <nc:group name="mdq:valueUnit">
                <nc:attribute name="@xlink:href" value="AnyValidURI"/>
            </nc:group>
        </nc:group>
        <nc:group name="mdq:result_d1e72">
            <nc:attribute name="__role" value="mdq:result"/>
            <nc:attribute name="__type" value="mdq:QE_CoverageResult"/>
            <!--
```

```
                             gmi:QE_CoverageResult:

                             The gmi:QE_CoverageResult is used to provide
a quality report that is in the form of a coverage (e.g. a
                             timeseries or grid). It therefore shares a
number of elements with coverage descriptions. In cases where
                             these elements are repeated in a single
metadata record, these can be references to elements that are
                             completely described elsewhere in the record
(e.g. xlink:href="#datasetspatialRepresentation").

                             The resultFile and resultFormat can be used
to describe a coverage result that is in a different file.
                        -->
            <nc:group name="mdq:spatialRepresentationType">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="mdq:resultSpatialRepresentation">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="mdq:resultContentDescription">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="mdq:resultFormat">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="mdq:resultFile">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
        </nc:group>
        <nc:group name="mdq:result_d1e80">
            <nc:attribute name="__role" value="mdq:result"/>
            <nc:attribute name="__type"
value="mdq:DQ_ConformanceResult"/>
            <!--
                             gmd:DQ_ConformanceResult:

                             The gmd:DQ_ConformanceResult is used to
provide a quality report on conformance of a dataset to a
                             particular standard, e.g. an OGC Web Service
or OPeNDAP.

                             The result in this case is a Boolean, i.e.
True or False.
                        -->
            <nc:group name="mdq:specification">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="mdq:explanation">
                <nc:attribute name="gco:nilReason" value="unknown"/>
            </nc:group>
            <nc:group name="mdq:pass">
                <nc:attribute name="gco:nilReason" value="unknown"/>
```

```xml
                </nc:group>
            </nc:group>
            <nc:group name="mdq:result_d1e86">
                <nc:attribute name="__role" value="mdq:result"/>
                <nc:attribute name="__type"
value="mdq:DQ_DescriptiveResult"/>
                <!--
                        DQ_DescriptiveResult
                        The DQ_DescriptiveResult provides an opportunity to
include a character string describing the quality test.
                        These are similar to the dataQuality element in the
DIF metadata standard.
                -->
                <nc:group name="mdq:statement">
                    <nc:attribute name="gco:nilReason" value="unknown"/>
                </nc:group>
            </nc:group>
        </nc:group>
        <nc:group name="namespace_1">
            <nc:attribute name="prefix" value="xml"/>
            <nc:attribute name="uri"
value="http://www.w3.org/XML/1998/namespace"/>
        </nc:group>
        <nc:group name="namespace_2">
            <nc:attribute name="prefix" value="xsi"/>
            <nc:attribute name="uri"
value="http://www.w3.org/2001/XMLSchema-instance"/>
        </nc:group>
        <nc:group name="namespace_3">
            <nc:attribute name="prefix" value="mdq"/>
            <nc:attribute name="uri"
value="http://standards.iso.org/iso/19157/-2/mdq/1.0"/>
        </nc:group>
        <nc:group name="namespace_4">
            <nc:attribute name="prefix" value="mdb"/>
            <nc:attribute name="uri"
value="http://standards.iso.org/iso/19115/-3/mdb/1.0"/>
        </nc:group>
        <nc:group name="namespace_5">
            <nc:attribute name="prefix" value="gco"/>
            <nc:attribute name="uri"
value="http://standards.iso.org/iso/19115/-3/gco/1.0"/>
        </nc:group>
        <nc:group name="namespace_6">
            <nc:attribute name="prefix" value="mcc"/>
            <nc:attribute name="uri"
value="http://standards.iso.org/iso/19115/-3/mcc/1.0"/>
        </nc:group>
        <nc:group name="namespace_7">
            <nc:attribute name="prefix" value="xlink"/>
            <nc:attribute name="uri" value="http://www.w3.org/1999/xlink"/>
        </nc:group>
        <nc:group name="namespace_8">
```

```xml
<nc:attribute name="prefix" value="cit"/>
<nc:attribute name="uri"
value="http://standards.iso.org/iso/19115/-3/cit/1.0"/>
```

```
            </nc:group>
        </nc:group>
    </nc:netcdf>
```

Notes

Tracking Compliance With Standards

The importance of standards in data integration is well know and broadly acknowledged. It is also well known that compliance tests will be required so that developers can systematically measure compliance with standards that evolve with time. Of course, these tests can only useful to users if the results are available and understandable.

There are two ISO Standard objects that apply to compliance reports. The first is the DQ_DataQuality object. This object associates a DQ_Element with a service, a dataset, or some subset of a dataset described by the DQ_Scope. The DQ_Element references a test and an evaluation procedure that produces a DQ_Result at a particular time. There are three types of DQ_Results, but the DQ_ConformanceResult is the type that is relevant here. It references the specification being tested, explains the meaning of conformance, and gives a Boolean result.

How might these new types of reports be used? The environmental data community is using more and more standard formats for data files. The conformance test might indicate how well the dataset conforms to those file formats or associated conventions like the Climate Forecast conventions for netCDF.

Tracking Problems Identified by Users

The second ISO object that could be related to conformance testing is the MD_Usage object. This provides a mechanism for incorporating user problems into the metadata for a dataset or service. The object includes a description of the specific usage, the problem encountered (the limitation), the date and the user contact information. In the case of standards compliance, the usage would be attempting to access the dataset or service in a way that is consistent with an advertised standard and the limitation would be finding the dataset to be non-compliant with the advertisement. These problem reports would be available from the archive and, hopefully, would motivate data providers to be careful about advertising services that they could not support.

MD_Usage
+ specificUsage : CharacterString
+ usageDateTime [0..1] : DateTime
+ userDeterminedLimitations [0..1] : CharacterString
+ userContactInfo [1..*] : CI_ResponsibleParty